

Revision memo

‘Do TJ policies cause backlash? Evidence from street name changes in Spain’, submitted to Research & Politics (Manuscript ID RAP-21-0006)

June 15, 2021

Response to Editor

Dear Editor,

We would like to thank you for the opportunity to revise the manuscript. As you will see, we have made thorough revisions, seriously considering each individual point raised by the reviewers. Below we address each of the reviewers in turn.

In general terms, we have XXXX Finally, we have also improved the empirical analyses, addressing all of the reviewers’ points on the analyses. We hope you will find that the manuscript has significantly improved through this process.

Response to Reviewer 1

We would like to thank the reviewer for all the comments. They were all very thoughtful and we found them very helpful in improving the manuscript. We respond to each of them below.

Comment 1: *On a conceptual level, I think the author(s) need to 1) define TJ policies more concretely, and 2) explain how renaming streets fits into the concept of TJ generally. The current definition of TJ offered in the manuscript (bottom of p2) combines truth-seeking, justice, and the politics of memory, but the case itself is exclusively about the politics of memory. The theory/discussion of contributions should reflect this focus on the politics of memory specifically, and highlight how these might be distinct from other types of TJ policies. In short, scope conditions need to be clearer.*

Response: Thank you for these suggestions. We have made an effort to clarify the scope conditions of our study. On the reviewer point 1) (definition of TJ policies), we have defined TJ policies more concretely, but we are a bit constrained by the word limit and that is why we refer to previous work on TJ, such as research by Jon Elster, Monika Nalepa, and Paloma Aguilar, among others. On 2), we clarify in the paper that street renaming, just like the removal of statues and other symbols or the building of museums and establishment of historical markers, are a form of symbolic Transitional Justice (Ward, 2021). In the case of Spain, as we detail in the paper, the renaming policy has been driven by the different TJ policy initiatives, particularly since the 2007 Law. Thus, we believe that this is one subset of TJ policies, and not just politics of memory, although these two things are very much intertwined. The reason why we keep our discussion within the discussion of TJ policies, and not exclusively on the politics of memory, is because we aim to make a contribution to the debate within TJ research on the effectiveness of these policies (following recent research by Capoccia and Pop-Eleches (2020) and Balcells, Palanza and Voytas (2020)). In any case, we are aware that the results in this study might not generalize to other types of TJ, and we now make sure that this is clearly stated in the manuscript.

Comment 2: *As it stands, I think the theory needs to be developed. Why do certain individuals react against TJ policies? Also: is asymmetric polarization context-specific to Spain, or a more general phenomenon?*

Response: We have revised the theoretical part of the manuscript substantially, and we have developed the theory further (within the word count constraints). Importantly, we have clarified that we do not expect an automatic asymmetric polarization effect of

TJ policies, and have explained that the process is as follows: 1) TJ policies can generate grudges among people who feel ideologically close to these symbols; 2) these grudges can be mobilized and politically exploited by political parties; 3) the latter can have political consequences, such as an increase in support for more extremist parties, if these are the parties mobilizing these issues; 4) all of this leads, in the case of Spain, to asymmetric polarization and increased support for the far-right. While grudges related to TJ policies will be different in other contexts (such as post-communism), we can see similar dynamics of political parties mobilizing and trying to capitalize these grudges. In this regard, this phenomenon can travel beyond the Spanish case. For example, there is recent evidence from Ukraine (Rozenas and Vlasenko, 2021) pointing to similar dynamics and outcomes in this post-communist country. In the case of Ukraine, asymmetric polarization does not occur at the right; it occurs at the “left”, with pro-Soviet parties gaining electoral ground after soviet symbols’ removals.

Comment 3: *On the broader contribution: I think there are two potential interpretations. Is backlash against TJ policies representative of the failure of TJ (ie: has renaming streets actually harmed the TJ process)? Or is backlash a second-order effect of TJ (ie: is successfully removing street signs a TJ victory with unfortunate side effects)?*

Response: We believe that backlash is a second-order effect of TJ, and not representative of the failure of TJ. After all, many Francoist symbols have been removed. In the long-run, we can expect that the removal of symbols of an authoritarian past will have other positive effects, such as a delegitimization of these regimes. Here, we are only looking at the short-term effects, and we cannot speak about long-term effects. We make sure that this is clear in the paper now. Also, we also clarify in the paper that, from a normative point of view, we are not trying to argue that these measures are not necessary. Even though they backlash, they might be necessary for healing and reconciliation purposes (Walsh, 2020; Ward, 2021). We believe that knowing what the unintended effects of TJ policies are, and precisely measuring them, is important. This also has relevant policy implications. For example, knowing that this is a phenomenon, traditional parties might get more ready to counter the extreme parties narratives around these issues, and minimize their potential backlash effects.

Comment 4: *On the empirics, my primary concern is that I can't fully judge whether the DiD is appropriate without more info. The unit of analysis in both analyses is the municipality; the DiD compares municipalities that changed street names to those that left them unchanged. However, the data section provides descriptive stats on street names rather than municipalities. Figures 1 and 2 show the total share of streets names in Spain that are Francoist and the number of individual street name changes, neither of which are relevant to the identification strategy*

per se. What I want to see:

- 1) The total number of municipalities that have/don't have Francoist street names in 2016 (sample size);*
- 2) The number of municipalities with Francoist street names in 2016 that changed one or more of these street names between 2016 and 2019 (ie: specify how many municipalities in the T and C conditions);*
- 3) How many street names were removed on average – and proportional to the total Francoist street names in the municipality – for municipalities that had street names removed (treatment strength);*
- 4) Some discussion of how municipalities in T/C conditions compare (ie discussion of Table 1 in Appendix, or just descriptive stats);*
- 5) Whether the trends in Vox/PP/PSOE vote share are parallel leading up to 2016.*

We completely agree with the reviewer that more information is needed to judge the identification strategy. We now included more additional information and graphs, both in the main text and in the appendix. In particular:

With regards to 1) and 2), we now include a frequency tables in the main text specifying the number of municipalities included in the DiD sample (i.e. municipalities that had Francoist street names in June 2016) and how many of these were in the treatment (street name removal between June 2016 and December 2018) and in the control groups. Moreover, as we explain below to other questions, we have significantly rearranged the Empirics section, and now our research design should be much more clear.

Regarding 3), we have included in the appendix a section called 'DiD sample and treatment strength' where we discuss the number of Francoist street name removals in treated municipalities, relative to the number of Francoist street they had at the beginning of the period, and how many Francoist street names they had not been yet removed at the end of the period.

Regarding 4), we include a more thorough discussion of Table 1 in the appendix, discussing the differences between treated and control municipalities in the sample (section A5 in appendix), along with many other analyses.

Finally, with regards to 5), we include in the main text a graph showing the parallel trends for the main parties of the analyses (PP and Vox, following also the recommen-

dation in the next comment), and we include much more detailed tables and analyses in the appendix (including test of the parallel trends in alternative DiD models for PP, Vox and PSOE with data since the early 2000s), in section A7.

Comment 5: *Also, on parallel trends:*

- 1) *I'd like a plot in the main text showing trends for PP and Vox aren't diverging in municipalities that removed names (or converging in non-removing municipalities) prior to 2016 – ie the potential endogeneity issue discussed at the bottom of p13. This is the key identifying assumption, and presenting tests (such as results from Table 5-6 from appendix, or just trend plots) in the main text is standard.*
- 2) *Author(s) test the assumption in Appendix Tables 5-6 but not for PSOE. Why not?*
- 3) *Vox only comes into existence in 2013, so there is only one election to test the assumption properly. Is the test for parallel trends in Table 5 still valid? Is there anyway to go further back? Some discussion of this is necessary.*

We appreciate this comment. We agree with the reviewer that showing that there are no diverging trends between control and treated municipalities is key for the credibility of the results. With regards to 1), we have now included a plot in the main text showing these parallel trends for PP and Vox, and include much more detailed analyses in the appendix (robustness tests for PP, PSOE, and Vox), in section A7.

With regards to 2), the reviewer is right in saying there is no reason not to include these robustness tests for PSOE. We did not include them in the beginning for reasons of brevity, but we now include a full table repeating all the same robustness tests for the DiD model (including parallel trends), but using PSOE as dependent variable. As shown in section A7 in the appendix, we now include data on elections since March 2000, when PP won its first absolute majority in parliament, thus testing parallel trends across different electoral periods that include right-dominated and left-dominated two-party systems and the emergence of a multi-party system. We also include a reference to these analyses in the main text, even though we cannot discuss their details for reasons of space. Overall, we do not find evidence of significant differences between municipalities in treated and control conditions, in terms of Vox (in 2015), PP (2000-2015) or PSOE (2000-2015). If anything, some models suggests that treated municipalities had shown stronger support for the PSOE in 2011 and 2015, although this result is not robust across all specifications. Moreover, we think that this should not be a concern for the analyses, as both our argument and results focus on a process of asymmetric polarization that mainly affects right-wing voters.

Finally (3), the reviewer raises a potential concern about the lack of pre-treatment data for Vox. This is a very relevant point. However, there is no easy solution to this problem as Vox did not participate in national elections before that year. Yet, we think that given that most electoral supporters of Vox had been PP voters in the past (e.g. [Rama, Zanotti, Turnbull-Dugarte and Santana, 2021](#)), testing the parallel trend assumption using PP votes (which holds for every year since 2000) should help to alleviate this potential problem.

Comment 6: *On the modeling:*

- 1) *Why do sample sizes in the DiD differ dramatically from outcome to outcome in Table 2? N should not change across outcomes, unless 1) vote shares are much more likely to be missing for Vox or 2) these are different samples – both of which would be a concern.*
- 2) *I'm not sure the cross-sectional analysis adds to the empirics, for the selection reasons the author(s) acknowledge. Maybe a triple difference with a second control category of municipalities that do not have Francoist street names in 2016 would similarly speak to generalizability of effects while retaining the nice identification strategy.*

With regards to 1), we agree with the reviewer that different sample sizes are a potential concern. The reason for these differences is that Vox did not participate in all municipalities in 2016, and that is why the sample is slightly smaller in this case. Yet, we agree that we should estimate the models on the same sample, and even though the results do not change, we now limit the samples in the PP and PSOE models to those municipalities included in the Vox analyses (we explain this in the main text). We also included a table (table 1, page 8) summarizing the DiD sample, so it can now be clearly seen.

Regarding 2), we agree that the cross-sectional models do not add much to the analyses, a point that was also raised by R2. Because of this, we have moved them to the Appendix (and now include different analyses with different specifications for the independent and dependent variables) and focus in the main text on the DiD analyses. We hope these changes also make the text much more clean and the research design easier to understand.

About the triple DiD models, we thank the reviewer for suggesting this possibility, which we were not aware of. We have extensively look into this method and, although interesting and with a lot of potential, we think it does not fit this research design and the data we have. As far as we have understood the difference-in-difference-in-differences (DDD), this method is meant to include a second control category. How-

ever, this second control group needs to be linked to the treatment group in some way that excludes the original control group. A classical example is a DiD design which compares the effect of a policy targeted at the elderly on the 55-64 (control) and 65+ (treated) population within a state (see [Wooldridge, 2010](#), 150–151), and expands the model into a DDD by introducing as a second control the elderly population (65+) of another state.

In our case, however, there is no way to separate statistically the second control (municipalities without Francoist street names in June 2016) from the treated population but excluding the original control (municipalities with streets in June 2016 but that did not change them during 2016–2018). Because of this, the DDD estimate (the triple interaction term, or $treated \times controlgroup1 \times t_1$) in a DDD equation with these three groups fails to estimate.

That said, if we have misunderstood this method and there is indeed a way to apply it to this design, we would be very grateful about some references on how to proceed. We think that this is a very interesting suggestion and that this estimation could significantly improve the analyses.

Comment 7: *Small things:*

- *The author(s) discuss Vox and PSOE, but (how) did PP politicize the issue?*

Response: We explain in the manuscript that PP, the mainstream conservative party in Spain, has been traditionally opposed to these policies, and they have been fierce advocates of the “pact of forgetting” that characterized the transition to democracy in Spain ([Aguilar, 2002](#)). However, Vox has been much more actively unapologetic about the Francoist regime, and has conducted a very aggressive whitewashing campaign, particularly in reaction of TJ policies such as street renaming. There has been an outbidding dynamic in the Spanish right in recent years, CITA which has benefited the far-right (Vox) at the expense of more moderate right-wing parties such as the PP or Ciudadanos. This outbidding dynamic has focused on the Catalonia secessionist conflict and historical memory issues.

- *Does the federal government coerce non-changers or reward changers? Might certain municipalities be more/less responsive to incentives?*

Response: We appreciate this very pertinent question. The short answer is no, the central government does not provide specific incentives to local governments for them to rename streets. As we mention in the main text, since these changes are mandatory following the 2007 Law of Historical Memory (LHM). The only ‘institutional coercion’

to change and remove the name of Francoist streets (along with other less common symbols) comes from the Ministry of Justice. For instance, in early 2019, the Ministry of Justice sent a letter to hundreds of local governments throughout the country, asking them to remove remaining Francoist street names (RTVE, 2019). In any case, the only effective mandate to remove streets comes from judicial sentences.

- *The author(s) say that the DiD sample is “probably” more rightist than the overall population of municipalities (p9). This can be tested empirically, and would be interesting.*

Response: We thank the reviewer for this very helpful comment. We now include a much more thorough comparison of municipalities in and out of the sample (including t-test comparison on electoral shares and logistic regressions) in the appendix (section A5). Also related to more comments by the reviewer about selection bias above, this is a very important point for the analyses. What we find is that indeed the municipalities included in the analyses are slightly more rightist than the rest, particularly when we only compare municipalities within the same region. We also include a reference in the main text to this section in the appendix.

- *I’d like to see a version of the DiD without controls.*

Response: We now include in the table on the main DiD models (table 3), models without control for each of the parties (Vox, PP, PSOE). Results do not change significantly.

- *Why do the author(s) simulate DiD estimates in Figure 3?*

Response: We use simulations to obtain an estimate of uncertainty of the DiD estimates. We draw 1000 set of coefficients from a multivariate distribution, and based on those draws, we simulate data. We then extract the DiD estimate for each simulation, as $(y_{treated,t1} - y_{control,t1}) - (y_{treated,t0} - y_{control,t0})$. Figure 3 shows the mean and 0.025 and 0.975 percentiles of these simulations.

- *Are the standard errors clustered on Autonomous Communities?*

Response: We do not use clustered standard errors on Autonomous Communities because the treatment is not taking place at the level of Autonomous Communities, and we already include fixed effects to account for potentially unobservables driving regional variation. That said, we agree with the reviewer that it is important to account for potential biases in the SEs, and now include a table in the robustness section of the appendix (section A7) with the results of the main models using heteroskedasticity-consistent SEs and SEs clustered at the level of municipalities. Although in some cases

the level of significance decreases a bit, and in some cases it increases, all key coefficients retain some level of statistical significance.

- *Figure 2: title says “percentage”, y-axis says “share”.*

Response: We have now changed the figure caption so they are coherent.

- *Possible typo on first paragraph of page 14: “dependent variable in continuous form (logged number of street name removals)”. Should this be independent variable? Also: how do the author(s) run a DiD with a continuous treatment?*

That is right, it is a typo. Thank you for pointing it out. We have now changed it to “the main independent variable”.

About running a DiD with a continuous treatment, we do in the conventional way, including an interaction between the time indicator and the treatment variable (which indicates treatment strength). The only difference is that the DiD estimate should not be interpreted in a binary way (the treatment effect) but as the effect of an increase in the (logged) number of Francoist street name removals on the treated. The other two coefficients are interpreted the same way.

Response to Reviewer 2

We would like to thank the reviewer for all the comments. They were all very thoughtful and we found them very helpful in improving the manuscript. We respond to each of them below.

Comment 1: *The literature Review and theory could be a bit more focused particularly to provide justification for the backlash idea. At the moment the expectation/hypothesis follows a very general literature review on the effects of transitional justice measures. More focus on the effect of political legacies on political support in the relation to transitional justice would help here as well. I would recommend looking beyond the literature on transitional justice and considering the growing literature on authoritarian legacies and political behavior in new democracies (see for example the Comparative Political Studies Special Issue Neundorff, A. and Pop-Eleches, G., 2020. Dictators and Their Subjects: Authoritarian Attitudinal Effects and Legacies. Comparative Political Studies, 53(12), pp.1839-1860.)*

Response: We thank the reviewer for this comment. In light of this comment, as well as R1's comments on our theoretical framework, we have revised the theoretical part of the manuscript substantially, and we have developed the theory further (within the word count constraints). First, we have clarified what we mean by TJ policies, and why we focus on "symbolic TJ" in particular. Second, we have developed further why we focus on the "backlash" hypothesis, building on early and more recent work on TJ justice (Balcells et al., 2020; Capoccia and Pop-Eleches, 2020; Elster, 2004). Third, in light of R2's thoughtful suggestion, we have more clearly referred to previous work on the legacies of authoritarianism in Spain, which has linked these legacies to attitudes towards TJ policies (Aguilar, 2002; Aguilar, Balcells and Cebolla-Boado, 2011), and political ideology (Balcells, 2012). Due to word constraints, we cannot address all the works in the CPS issue they refer us to, unfortunately, but we do refer to some of Pop-Eleches's work on this topic.

Comment 2: *The initial motivation is a bit confusing by focusing on the US and then switching to Spain. Why the US? Surely taking down symbols of the old regime is a more general phenomenon? Also the US perhaps does not fit the more general paradigm of a post-authoritarian context. And why is Spain a good case to examine the effects of taking down symbols? Are the mechanisms proposed to be at work in Spain generalizable to other contexts? Why do you assume that taking down symbols of the old regime will necessarily increase support for the far-right? Why not extremist parties in general? See Milan Svolik's paper (Svolik, M.W., 2019. Polarization versus democracy. Journal of Democracy, 30(3), pp.20-32.) How would what you find be generalized to post-communist contexts? Do you think taking down communist*

symbols should lead to a revival of some form of left-wing parties?

Response: We appreciate these thoughtful inquiries. We have now downplayed the US example in the introduction. However, we believe that, in spite of the differences between the US and Spain, the US parallel works because historical memory symbolic politics matter a great deal for contemporary politics too. The US is not a paradigm of a post-authoritarian context, but the Confederacy was a repressive regime in which a subset of the population was enslaved, systematically discriminated, and oppressed. Thus, removing the symbols of the Confederacy is a way to somewhat acknowledge these victims, and try to achieve some symbolic reparations, which can help in the ultimate goal of achieving reconciliation (Walsh, 2020; Ward, 2021). As Grossman puts it, “Choices are made about what gets built, displayed, and given plaques. Memorials are public commemorations that legitimate what comes to be called ‘heritage’” (Grossman, 2016).

With regards to the theoretical mechanisms and the external validity of the theory, we have clarified that we do not expect an automatic asymmetric polarization effect of TJ policies everywhere, and have explained that the process is as follows: 1) TJ policies can generate grudges among people who feel ideologically close to these symbols; 2) these grudges can be mobilized and politically exploited by political parties; 3) the latter can have political consequences, such as an increase in support for more extremist parties, if these are the parties mobilizing these issues; 4) all of this leads, in the case of Spain, to asymmetric polarization and increased support for the far-right.

We believe that this theory could travel to other contexts, although with context-dependent considerations. For example, regarding the reviewer’s question about post-communist contexts, we would expect that taking down communist symbols would cause increased support for parties associated with the communist past if they mobilized and tried to exploit this issue. In a recent working paper we now cite, Rozenas and Vlasenko (2021) find that the removal of Soviet monuments in Ukraine (during the so-called *Leninopad*) mobilized supporters of the Soviet legacy parties, which, they argue, “are those with a relatively favorable interpretation of Ukraine’s Soviet past.” These are the parties that exploit popular discontentment against these removals and provoke a political backlash. Again, we now clarify in the manuscript that we do not expect the removal of the symbols to lead to an automatic increase of support for specific parties: this will happen only if such parties are mobilizing and trying to electorally exploit grudges around these removals. In the case of Spain, these parties are the far-right parties, which have connections to the Francoist past, and openly vindicate it.

We now also explain more clearly why Spain is a good case to analyze the removal of street names, which, just like the removal of statues and other symbols or the building of museums and historical markers, are a form of symbolic Transitional Justice (Balcells et al., 2020; Ward, 2021). In Spain, TJ took place several years after the end of the dictatorship and transition to democracy (Aguilar, 2002). Spaniards are quite divided on their opinions regarding TJ measures, and these divisions roughly fall into the left-right divide, with the left being more favorable to them than the right (Aguilar et al., 2011). The removal of Francoist street names is a TJ measure that, together with others such as the removal of General Francisco Franco from its mausoleum (“Valle de los Caídos”), has been challenged by right-wing parties, and particularly so by the extreme-right. It is thus a quite ideal case to assess if these measures backlash and electorally benefit these parties.

Comment 3: *Also how generalizable are the findings even within Spain? Is the backlash specific to areas that might have high latent rightist support? This issue is also linked to concerns about the DiD design.*

Response: This is a great comment, and as the reviewer discusses further down, the external validity of the results might not be as strong. We think there are two answers to this point.

On the one hand, in terms of aggregate processes, it would indeed be surprising to observe a similar backlash in regions such as Catalonia or the Basque Country, where support for Spanish nationalist parties is overall slim, and where most Francoist symbols were removed shortly after the transition to democracy. So in this sense, the findings—in terms of *how much* support for Vox was attributed to the removal of symbols—would not travel so easily.

On the other hand, however, the answer above suffers from an ecological fallacy. It is true that the level of ‘Francoism sympathizers’ is higher in the municipalities we study than in the rest, but that does not mean that even in the most staunchly leftist areas there are individuals who have a positive view of the Francoist past. And what we are ultimately studying here is an individual response, even if mediated by the supply side—i.e. mobilization strategies of political parties—or the social context. Thus, we think that the theory and the mechanism we identifies can also be applied in other regions, even if the overall impact is smaller.

Actually, anecdotal evidence supports this last point. For example, in Tortosa, in Catalonia, people voted in 2016 against the removal of a very prominent Francoist symbol. This included the mayor of the town and the councillors for CiU, the Catalan national-

ist party (Berbís, 2016). The symbol was only removed after a judicial sentence.

Comment 4: *The methodological and empirical sections need to be better organized and more transparent. The relevant information for understanding the DiD design is scattered throughout the methodology section and the first part of the analysis section. The word count limit is low, but specific information on the measurement levels of each variable would be helpful as well as the model equations. For the robustness checks referred to in the appendix, please provide specific table numbers in the appendix in the main body of the text. It would be clearer to introduce the variables, then run the main analysis, and then refer to the robustness checks in the appendix. As an example the author(s) chose to code the Francoist street name removals as a binary variable. This is an important decision. I would like to see some graphical representation of how many and which municipalities would fall into either category. For the cross-sectional analysis I would be interested in seeing the effects of the Francoist street name removals as an interval measure in relation to the effects of the binary coding. The authors say they have done the additional analysis using an interval measure (logged number of changes) but only in the methods section (page 6) and also it is not very clear which table they refer to in the appendix. It would be logical to mention this robustness check after the discussion of the results in Table 1 of the main body. It is not clear which table shows the cross-sectional analysis with the logged number of removals.*

Response: We are very thankful to the reviewer for this detailed comment on the empirical analyses. R1 also raised some comments about the Empirics section. We have significantly rearranged this section so we hope it is now much cleared and easier to understand. In particular:

We have removed the cross-sectional analyses from the main text and placed them in the appendix, as both R1 and R2 suggested that their added value was limited. This has allowed us to focus on the DiD analyses and better explain them. In addition, the cross-sectional analyses in the appendix now include models with both binary and continuous versions of the independent variable.

We follow the suggestion by the reviewer and we now 1) present the research design and the variables, 2) run the main analyses, and 3) explain the robustness checks and alternative models, with explicit references to sections in the appendix.

We have included much more information on the variables, even though we could not explain it at length in the main text because of space constraints (we do include a table in the main text specifying how many municipalities fall into the sample criteria and how many of these fall in the treatment and control groups, table 1).

We now include in the appendix summary statistics of the variables (section A2), a map showing the municipalities included in the DiD analyses (section A2, including the full sample and the limited sample excluding municipalities where Vox did not participate in 2016 elections), information of treatment strength (i.e. what is actually behind the binary coding, showing that most municipalities had indeed few streets to begin with and few changes) (section A3), and a very detailed comparison of municipalities in and out of the sample and treated vs. control municipalities (section A5).

Comment 5: *I would also generally be interested in understanding more about what distinguishes the municipalities that retained Francoist street names until the periods analyzed from those that removed the names before 2001. Perhaps a map of Spain showing where are the municipalities included in the sample are located would be nice. What is their history? Are they former Francoist strongholds? If so the effects of removal of symbols would be rather inflated: that is if we could go back in time and look at the earlier effects of the removals of street names would they have caused an equally sized backlash? Maybe not if the removals were done in areas of Spain that were more anti-Franco.*

Response: We agree with the reviewer that this is a major point affecting the interpretation of the results, which was also pointed out by R1. We now include a much more detailed section in the appendix (section A5) making a thorough comparison between those municipalities in the sample and those that are not in it. We mention it in the main text but, beyond what we say about municipalities that still retained Francoist street names in 2016, space constraints prevent us from a more detailed discussion. We do hope that all the information in the appendix helps the reader to evaluate our discussion of any potential selection biases.

Regarding the results of these analyses, as we mention in the main text, we find that municipalities in the sample are slightly more rightist than those outside it. In terms of the interpretation of the results, we think it does not harm our inferences. One could say that a more rightist sample makes the generalization of the results more difficult, as the same effect would not have taken place in more leftist areas. However, as we discuss above in Comment 3 of R2, we think this reasoning could be an ecological fallacy. The fact that a municipality shows higher support for rightist parties does not mean that everybody is rightist, nor that in a more leftist municipality everyone is leftist. Thus, we do think that the early removal of Francoist symbols in areas where support for Spanish nationalism was much lower, such as Catalonia or the Basque Country, also caused a backlash effect in some individuals, even if the aggregate effect was much more limited because of a lower number of rightist sympathizers.

In terms of our analyses, we think that, if anything, more rightist municipalities should already be more likely to support the new far right party Vox, which should in some way cancel out the effect of street name removal. A more leftist sample would be more problematic as we are dealing with a potential backlash effect among the rightist population in areas where right-wing political parties are less likely to enjoy local power. Obviously, a more detailed analyses using individual data, if that data were available, would help to dig deeper into these questions.

Comment 6: *For the first cross-sectional analysis: what is the point of the model (Table 1) on the full sample if the majority of municipalities had no Francoist street names to remove? In other words these municipalities are not eligible for “treatment” anyhow and more likely to be fundamentally different from the municipalities that still retained Francoist street names in 2001 or later. For the cross-sectional analysis I would be interested to see the effects of the interval/continuous measure of Francoist street name removals. I would also like to systematically see maybe in the appendix side by side the regression results using varied periods in which the Francoist street name removals are considered (e.g. 2001-2018, 2011-2018, etc). The main point is that we have some systematic change in the samples so for example same end point to the period but sliding starting points. I have also noticed that the samples vary according to whether all the included municipalities had some Francoist street names in the starting year of the period considered for the independent variable. It is logical obviously to look at the municipalities that had Francoist street names at the starting but again consistency would help (Table 4 in the appendix states the sample only includes municipalities that had street names in June 2011 but the period considered for removal is December 2010-2018).*

Response: We appreciate this comment, which also resonates with some of the points raised by R1 about the relevance of the cross-sectional analyses. In response, we have moved the cross-sectional analyses to the appendix, where we test many of the suggestions above (section A6), and focus on the DiD models in the main text.

In addition, we think the reviewer is right in saying that the first two models of the cross-sectional analysis do not add much to the paper, we have now dropped them from main analyses, and included instead two models using the main independent variable (the number of Francoist street name removals) in continuous form, as per the suggestion of the referee. Results do not significantly differ.

Moreover, we also include in the appendix (section A6 as well) a table with different cross-sectional models using the sliding starting point that the reviewer suggests. Coherent with the main findings of the paper, the results show that it is only more recent street name changes that correlate with higher support for Vox.

Moreover, we have corrected the table (previously table 4 in appendix, now table 10). The reviewer is right in saying that the tables needed consistency, so we have now corrected all notes and mentions and always talk about June or December to refer to the periods. The reason of the previous inconsistencies was that the actual dates we have data on street names is always December 31st and June 30th, that why we sometimes said that we limit the samples following streets in January. We hope is now much more clear in the main text, where we explicitly explain this.

Comment 7: *The application of the difference-in-difference design could be much clearer. The word count limit is quite short but briefly writing out the model (Also clearly specify which in the models coefficients capture the treatment effect) and perhaps explaining the design explicitly in terms of treatment and control would help (additional details could go in the appendix). Noting details like the model type would help too (I am assuming OLS based on the dependent variable). Overall it seems like the design should be viable and defensible although the generalizability of the findings might be more limited than currently stated. But I would like to see more explicit consideration of DiD design. To this end it would help write out the model and in the appendix maybe explain the more design more systematically in terms of treatment and control and expected outcomes. Telling us the mean values on the outcome for these municipalities versus the “treatment” municipalities as well as the levels of Francoist street names.*

Response: We thank the reviewer for all these suggestions aimed at clarifying the analyses. Following comment 4 above, we now include the model equation for the DiD model (as we have moved the cross-sectional analyses to the appendix) and mention which coefficient is capturing the treatment effects.

Regarding discussing the model in terms of treatment and control groups, we are a bit hesitant to do so extensively in the main text because, strictly speaking, we are not dealing with experimental data. However, we hope that after having rearranged the Empirics section is now much more clear. In any case, we do speak of treated and control municipalities when we present the sample in Table 1, and when we show base differences in shares between the groups in Table 2. In the appendix, we include a much more extensive discussion of selection bias both in terms of inclusion in the sample and treated vs control (section A5), where we do use the experimental terms more freely.

We how this helps to clarify our design without including false pretences about experimental designs. Along the same lines, and following several other comments raised by both R1 and R2, we also include much more evidence in the appendix on treatment strength, and balance between treated and control municipalities (see sections A3 and

A5 in particular).

Comment 8: *Next the author(s) briefly mention the issue of “selection bias” and how the “control” group municipalities could be more right-wing. I would like to know more about all this and how the design is still justifiable. Let us assume for example that the municipalities in the control group that have retained Francoist street names and resisted changing them did so because they are right-wing strong holds, how would this characteristic affect support for Vox? Is it possible that the support for Vox was established earlier before the period analyzed and reached a ceiling? So maybe we do not have much change/variation in Vox support to explain among the control group because the support appeared earlier and remained stable? I am also puzzled whether the level of name changes is the only likely trigger for the increase in Vox support. If there is so much potential for far-right support, who was in control of these municipalities and who drove through the street name removals? Was it left-wing local governments? Are we seeing perhaps the effects of municipal level polarization driven by the national discourse? The author(s) note that two events coincided: the removal of street names and Vox nationalist campaign. Could both be responses to overall political polarization? It would be also interesting to have more discussion about the finding that support for the more relatively more centrist PP declined. Again see debates related to Svolik 2020. Some of this issues might partially be addressed through interpretation of Table 1 in the appendix.*

CHECK

Response: We are grateful again for this inquiries. The reason we say the control group is more rightist is because we limit the sample to municipalities that still had Francoist streets in June 2016, so this comparison is between the control group and the municipalities out of the sample. We have added “(compared to municipalities out of the sample)” to the main text (page 8) to clarify this. Moreover, we offer several analyses in the appendix (section A6) where we test this assumption, showing comparisons both between sample vs out-of-sample and control vs treatment groups. Also, we have included a graph in the main text (figure 4) showing pre-treatment trends for PP and Vox.

That said, the reviewer is right to ask how this affects the results. If municipalities in the control group (municipalities with Francoist streets in June 2016 that did *not* change them in the 2016–2018 period) are indeed more rightist than treated municipalities (even if this is not reflected in the pre-treatment trends), we think this selection bias would go against our results. In other words, Vox would have become stronger in those municipalities, whereas our main finding is that it grew more in treated municipalities. So by limiting the sample to this relatively rightist municipalities, we get rid of all left-

leaning municipalities that could slid into the control group and bias our results.

Similarly, the reviewer is right when asking about who changed this streets if we should be in ‘right-wing territory’, and whether nation-wide or local trends of polarization could have affected both the change in street names and the increase in Vox support. The way we see this is that there are two potential explanations behind the change in street names in such a relatively late period. One could be a shift in local power to the left, which most likely was the reason in some municipalities (the city of Madrid, for instance). Yet if this was a more general patterns, local polarization could be driving the results, confounding the relationship between name removals and changes in electoral support. However, we find in the analyses on name changes (table 2 in appendix, section A5) that having a leftist mayor in 2015 is not significantly correlated with name changes. Also, we do not find any effect on vote share for PSOE. If there was generalized polarization, on both sides of the ideological spectrum, we should find a decrease in support for PSOE, similarly to what happens with PP.

Another possibility is that ‘late’ name changes was just the result of the implementation of the 2007 Law of Historical Memory which, unless local governments were active and supportive, did take some time to come into effect. Even if the government was not too insisting at the beginning, over time the outcome of some legal procedures reminded or forced many municipalities to make those changes. So what we are capturing is the effect of this changes on internal electoral shifts within the right-wing electorate, which became more polarized.

Regarding nation-level dynamics, such as the presence of nationalist discourses, they should affect all municipalities equally, and although there were definitely some interaction processes at the local level (perhaps some local politicians were more active in campaigning in this discourse), it is a dimension that, even if very interesting, falls outside of the scope of this paper. Yet, we do think this comment raises a very interesting point which could definitely motivate further work.

Comment 9: *Also on a more detailed technical note what about first differencing the DiD model? Then you could better leverage the variation in the treatment effect. See discussion of Card 1992 in the section on DiDs in Angrist, J.D. and Pischke, J.S., 2008. Mostly harmless econometrics: An empiricist’s companion. Princeton university press.*

Response: We thank the reviewer for this suggestion, which we now include in section A8 in the appendix, including first-difference models for the period of interest and the two immediately before and after (always including the removal of Francoist street names between June 2016 and December 2018 as independent variable). Results from

FD models are coherent with the analyses, showing an increase in support for Vox and a decreasing in support for PP only during the 2016–2019 period.

Comment 10: *I would also like a clearer and more explicit discussion of the parallel trends assumption in the appendix. It would help to see some visual representation of party support over time in the municipalities included in the DiD analysis. I would expect to see that that difference (could be no difference) in party support for the right-wing parties (PP, Vox) between in the “control” municipalities (those that had Francoist street names and did not change them) and those (that removed Francoist street names) remained stable before the “treatment period.” For the support for Vox it appears that there is only one time point (2015 election) before the pre- and post-treatment time periods. Really to establish a pre-treatment parallel trend more than one time point is needed ideally. Maybe the PP vote could be a proxy. We should expect the PP vote relative to the Vox vote to be stable in the control group and stable difference between PP votes in the treatment vs control going back in time.*

Response: We thank the reviewer for this insightful comment, which points at a key aspect of the analyses. We now include much more information on the parallel trends assumption in the appendix (section A7) and include as well a graph in the main text (figure 3) showing pre-treatment trends for PP and Vox. All the analyses show that there are no different in support for these two parties between control and treatment groups before the period of study. In the case of PP (and PSOE, included in the appendix) we should analyses of pre-treatment trends as far back as the March 2000 elections.

The reviewer is right to mention that only having one observations before the treatment for Vox could be problematic. However, it is impossible to look further back. He is also right in saying that PP vote could be a proxy, and indeed is, since most Vox supporters were PP voters in the past. So that is our strategy: to test the parallel trend assumption using vote share for PP. We also include PSOE for comparison, but given our findings, the most worrying result would be to find divergent trends within the rightist population.

Comment 11: *As to the conclusions and implications: why should the removal of old regime symbols just increase in support for the far-right? If the idea of backlash is to travel, then surely we would expect in Eastern Europe an increase in support for the left if communist symbols are removed?*

Response: As argued above, we expect the removal of symbols to generate backlash among supporters of the former regime or the perpetrators, and thus to increase support for parties that are mobilizing this issue and that try to electorally benefit from

these grudges. These are not necessarily far-right parties; they are in the case of Spain because the far-right is connected to the past authoritarian regime (i.e. Francoism), vindicates it, and it mobilizes the issue of the symbols. Traditional conservative parties are less adamant to defend such symbols, or they do it less fiercely, so the far-right ends up benefiting from it. In a way, there is an outbidding dynamic ([Kydd and Walter, 2002](#)) on this issue which benefits the most extreme parties. As discussed above, there is some evidence from a former communist country (i.e. Ukraine) that supports the idea that these backlash effects travel to other contexts, and that parties that defend the soviet legacy benefit from such backlash. Similarly, in the case of the US, we would expect that the removal of Confederate symbols would benefit electoral candidates that defend the legacy of the Confederacy and that fiercely oppose the removal. This is something that we could test in further research.

Again, we very much appreciate the opportunity to revise and resubmit our manuscript. We believe that these thoughtful comments have helped us considerably in strengthening the article. To recap, we we have XXX, we have clarified some aspects of our theory that were confusing, and we have revised the empirical tests to document our findings' robustness and further substantiate our claims. We made all these changes staying within the journal page limit. We hope that you are happy with these changes, and we are deeply grateful for your time and assistance on this project.

References

- Aguilar P (2002) *Memory and Amnesia: The Role of the Spanish Civil War in the Transition to Democracy*. New York: Berghahn Books.
- Aguilar P, Balcells L and Cebolla-Boado H (2011) Determinants of attitudes toward transitional justice: An empirical analysis of the spanish case. *Comparative Political Studies* 44(10): 1397–1430.
- Balcells L (2012) The consequences of victimization on political identities: Evidence from Spain. *Politics & Society* 40(3): 311–347.
- Balcells L, Palanza V and Voytas E (2020) Do museums promote reconciliation? a field experiment on transitional justice. *The Journal of Politics* forthcoming.
- Berbís S (2016) Tortosa vota este sábado si retira el monumento franquista del Ebre. *El Periódico*, May 28, 2016. Available at <https://bit.ly/3ws9UyB> (accessed 15/06/2021).
- Capoccia G and Pop-Eleches G (2020) Democracy and retribution: Transitional justice and regime support in postwar west germany. *Comparative Political Studies* 53(3–4): 399–433.
- Elster J (2004) *Closing the books: Transitional justice in historical perspective*. Cambridge University Press.
- Grossman J (2016) Whose memory? whose monuments? history, commemoration, and the struggle for an ethical past. *Perspectives on History* (February).
- Kydd A and Walter BF (2002) Sabotaging the peace: The politics of extremist violence. *International Organization* 56(02): 263–296.
- Rama J, Zanotti L, Turnbull-Dugarte SJ and Santana A (2021) *VOX: The rise of the Spanish populist radical right*. London: Routledge.
- Rozenas A and Vlasenko A (2021) The real consequences of symbolic politics: Breaking the soviet past in ukraine. SSRN Working Paper, last revised April 20, 2021. Available at <https://bit.ly/3gupfro> (accessed 11/06/2021).
- RTVE (2019) El ministerio de justicia pide a 656 ayuntamientos que retiren los vestigios franquistas de sus calles. RTVE.es / EFE, February 6, 2019. Available at <https://bit.ly/3gjuLyo> (accessed 14/06/2021).
- Walsh C (2020) Must we allow symbols of racism on public land? The Harvard Gazette, June 19, 2020. Available at <https://bit.ly/3xfS41K> (accessed 11/06/2021).

Ward G (2021) Monumental antiracism. a collaborative survey of anti-racist memory work. Updated May 21, 2021, available at <https://bit.ly/3gsg6zz> (accessed 11/06/2021).

Wooldridge JM (2010) *Econometric Analysis of Cross Section and Panel Data*. MIT press.